



## Spatial analyses to evaluate multi-crop yield stability for a field

J.M. McKinion\*, J.L. Willers, J.N. Jenkins

Genetics and Precision Agriculture Research Unit, P. O. Box 5367, USDA-ARS, Mississippi State, MS, United States

### ARTICLE INFO

#### Article history:

Received 15 January 2009

Received in revised form

28 September 2009

Accepted 7 October 2009

#### Keywords:

GIS

GPS

LIDAR

Precision agriculture

Yield monitor

Digital elevation map

Spatial analysis

Cluster analysis

Terrain features

### ABSTRACT

This paper proposes that yield stability patterns exist for multiple crops planted on the same land area over a period of years that growers can use to their advantage in planning crop management strategies using precision agriculture technologies. This study examines the relationship of soil elevation, slope, aspect and curvature to crop yield stability using a digital elevation model of the study area derived from a precise light detection and ranging (LIDAR) image of the farming area and surroundings. Three crop years of cotton and two crop years of corn yields were used to evaluate this hypothesis. The interpolation methods of Inverse Distance Weighted (IDW), simple Kriging and Natural Neighbor found in ESRI's ARCGIS were used to produce crop yield maps. These methods were also compared in the analysis. Simple Kriging gave the best  $R^2$  estimates of yield as a function of elevation, slope, curvature and aspect. When the SAS FastCluster procedure was used to group yield points together using topographical features, the resulting regression analyses  $R^2$  values of yield as a function of elevation, aspect, curvature and slope by cluster number were improved.

Published by Elsevier B.V.

### 1. Introduction

Precision agriculture has been defined as “applying the right practice at the right time in the right amount” and “farming by the square foot”. With the advent of geographical positioning systems (GPS) and microcomputer-based controllers on farm equipment, the application of precision agriculture technology has generated considerable amounts of data and data mining opportunities for researchers. Or put another way, “We are drowning in information and starving for knowledge – Rutherford B. Rogers” (Bullock et al., 2007). As each crop year finishes, growers collect yield maps of their fields as generated by yield monitors and GPS sensors on their harvest equipment. These maps are in addition to maps generated by variable rate applications involving pre-plant chemical applications, precision planting and post planting chemical applications, which can be extensive and numerous according to the crop planted and to the as-applied maps collected by the application equipment which can be different from the prescription.

In addition, some growers have automated weather stations on their farms and record temperature, humidity (or dew point), rainfall, solar radiation, and daily wind run. The challenge is to use these data to help growers better manage their crops. While traditional replicated small plot research has excelled in finding differences

among means for treatments, a new opportunity has arisen with the advent of precision agriculture. New methods are available that can assist growers in better management of their limited resources (Bullock and Bullock, 2000; Drysdale and Metternicht, 2003; Fleming et al., 2000; Hornung et al., 2006; Kaspar et al., 2003; Kravchenko and Bullock, 2002a,b; Kravchenko et al., 2005; Mueller et al., 2001).

According to Kaspar et al. (2003), the use of yield maps in decision making for the next season is difficult because of the problems of interpretation. There are many mitigating factors. Permanent spatial factors that affect yield either directly or indirectly are landscape position, terrain attributes, erosion class and soil properties (Spomer and Piest, 1982; Stone et al., 1985; Jones et al., 1989; Kravchenko and Bullock, 2000). Transient spatial factors which can affect yield in specific areas in one year but not every year are insects, disease pathogens and planter or applicator malfunctions. Measurement errors in the yield monitor equipment can also occur (Lark et al., 1997; Colvin et al., 1997). As a result errors that occur in one year can obscure patterns in the yield map making it difficult or impossible to discern clear trends or patterns (Kaspar et al., 2003). Thus, maps from several years are needed to discern these patterns.

Digital Elevation Models (DEM) and Remote Sensing data provide information about the earth's surface and can aid in determining characteristics of the landscape and the soil (Drysdale and Metternicht, 2003). In this research we hypothesized that terrain features of a field are significantly related to crop yield across years

\* Corresponding author. Tel.: +1 662 320 7449; fax: +1 662 320 7528.

E-mail address: [james.mckinion@ars.usda.gov](mailto:james.mckinion@ars.usda.gov) (J.M. McKinion).

and crop species. If this hypothesis is correct, knowledge of these terrain features can be used to improve management of the crop irrespective of crop species.

The specific objective of this research is to explore the hypothesis that yield stability patterns exist for multiple crops planted on the same land area over a period of years that growers can use to their advantage in planning crop management strategies using precision agriculture technologies. Further, high resolution LIDAR information on field surface topology can aid in the spatial analysis process, and the statistical analysis procedure cluster analysis can improve goodness of fit in applying regression analysis.

## 2. Spatial variability

Kravchenko and Bullock (2002a) reported that environmental factors affecting crop growth are often spatially variable and this variability extends from microscale to field and watershed scales. Soil characteristics are often distributed within fields in a manner that is amenable to geostatistical description and analysis (Warrick et al., 1986). The variability of environmental and soil factors contributes to variability in crop performance. Miller et al. (1988) observed spatial correlations between wheat harvest indices and soil clay content. Jaynes and Colvin (1997) observed spatial structure in within-field crop yield variability for 6 years of corn and soybean yields in Iowa.

Kravchenko and Bullock (2002a) state that topography is a particularly attractive variable in describing and predicting spatial variability of crop yields for precision agriculture management. It is a soil formation factor that defines distribution of soil moisture, organic matter content, nutrients, soil textural composition and soil physical properties which affect crop growth and yield within a field (Changere and LaL, 1997; Hanna et al., 1982; Moore et al., 1993). Topography also affects temperature and humidity variations within a growing crop. Thus, topography can be regarded as a compound parameter that reflects the combined influence of various yield-affecting factors and interactions. Topographical variables such as elevation, terrain slope and curvature can explain from 6 to 54% of the variability of corn and soybean yields (Kravchenko and Bullock, 2002a). Timlin et al. (1998) found surface curvature to be a useful parameter for describing yield, topography, and weather relationships.

## 3. Remote sensing

Optical remote sensing measurements record the radiation emitted and reflected from the soil surface (Drysdale and Metternicht, 2003). There is very little penetration of electromagnetic energy through the soil body. Soil reflectance derives from the inherent spectral behavior of the heterogeneous combination of the biochemical (mineral and organic) constituents, geometrical-optical scattering (particle size, aspect, roughness) and moisture conditions of the surface (Baumgardner et al., 1995; Ben-Dor et al., 1998; Irons et al., 1989).

The capability to create digital surface models (DSMs) of elevation for rural or urban landscapes is facilitated by light detection and ranging (LIDAR) or laser scanning sensor systems (Axelsson, 1999; Baltsavias, 1999; Jensen, 2000; Wehr and Lohr, 1999; Willers et al., in press). Digital surface models (DSMs) are topographic maps of the earth's surface that provide one with a geometrically correct reference frame over which other data layers can be draped (Intermap Technologies, 2009). For this study we use the generic term DSM for representing the topographic features of elevation, slope, aspect, and curvature except when explicitly defining a single component, i.e. DSM of elevation. A DSM describing landscape elevations creates opportunities for solving many problems (i.e.,

**Table 1**  
Lidar specifications.

|                     |            |
|---------------------|------------|
| Flying height       | 3000 m AMT |
| Airspeed            | 110 knots  |
| Laser pulse rate    | 43 kHz     |
| Field of view       | 25°        |
| Scan rate           | 40 Hz      |
| Average swath width | 405 m      |

information about vegetation, vegetation growing under power lines, automatic capture of buildings for modeling purposes, extraction of breaklines (road edges or hydrographics features) from terrain features) (Ackerman, 1999; Filin, 2004). DSMs of elevation have been widely applied in forestry (Kraus and Pfeifer, 1998; Means et al., 2000; Popescu et al., 2002), bare-earth extraction (Sithole and Vosselman, 2004), urban planning (Shan and Samarth, 2005) and many other applications (Barnes et al., 1990; Hollaus et al., 2005; Leyva et al., 2002). It is essential that LIDAR data be of high quality (Latypov, 2002) in all of these applications, particularly for agricultural fields that have low relief. Sources of errors in LIDAR data can be apportioned into (1) random and/or (2) systematic causes and (3) blunders (Baltsavias, 1999; Huising and Pereira, 1998) and/or random errors, corrected by manual or automated filtering methods (Früh and Kilian, 1994). Systematic errors are due to characteristics of the laser scanner itself and/or the internal navigation system (INS) of the aircraft carrying the scanner. Systematic errors must also be corrected in order to produce good quality DSMs of elevation (Morin, 2002; Skalud and Lichti, 2006; Vosselman, 2002).

## 4. Materials and methods

This study was conducted on the Paul Good Farm located in the northeast central edge of Mississippi in Noxubee County, MS, USA and consists of approximately 600 hectares of gently rolling contiguous farm area. The field used for this study was field F160 which consists of 65 hectares (160 acres) and is bounded on all four sides by farm roadways. Farming enterprises include dryland crop rotations of cotton, corn and soybeans.

## 5. Digital elevation model

EarthData Aviation<sup>1</sup> (Hagerstown, MD) acquired the data to generate the digital elevation model (DEM) for the Paul Good Farm on May 12, 2003 using its Navajo aircraft. A DEM is the representation of continuous elevation values over a topographic surface by a regular array of z-values, referenced to a common datum. DEMs are typically used to represent terrain relief (GIS Dictionary, 2009). LIDAR data was captured using an ALS40 LIDAR system, including an inertial measuring unit (IMU) and a dual frequency GPS receiver. The areas of interest were overflown at an altitude of 3000 m (9842 ft.) above the terrain. The LIDAR specifications are given in Table 1.

LIDAR, IMU, and GPS data were correlated using GPS time and processed using LIDAR post-processing software to determine the coordinate of each point on the ground. The AeroScan LIDAR is able to receive up to three returns from each laser shot fired. This allows receipt of return data from multiple objects as the laser beam travels towards the ground. All ranges have been corrected for atmospheric refraction and transmission delays. The resulting three dimensional coordinates are compiled in a mass point file of

<sup>1</sup> Mention of trade names or commercial products is solely for the purpose of providing specific information and does not imply recommendations or endorsement by the U.S. Department of Agriculture.

**Table 2**

Accuracy of the LIDAR data compared to kinematic ground data points for the Paul Good Farm on May 12, 2003.

|                            |              |
|----------------------------|--------------|
| Vertical accuracy          | 10.7 cm RMSE |
| Standard deviation         | 10.7 cm      |
| Mean difference            | 0.0 cm       |
| Number of points in sample | 481          |

x, y, z on the UTM projection. Ellipsoidal heights were converted to NAVD88 using Geoid 99. These files were in an EarthData binary format (E-EBN). Initial evaluation of the LIDAR data was accomplished including comparison of the data from flight line to flight line and against the walk-around, ground survey.

A walk-around survey by ground personnel was conducted by EarthData Aviation to collect profile data of taxiways and runways at the base airports, as well as within one of the project sites. These data were supplemented with several NGS control station points located within the project area. This survey was used for accuracy verification of the processed LIDAR data.

The comparison with the kinematic survey yielded the results shown in Table 2.

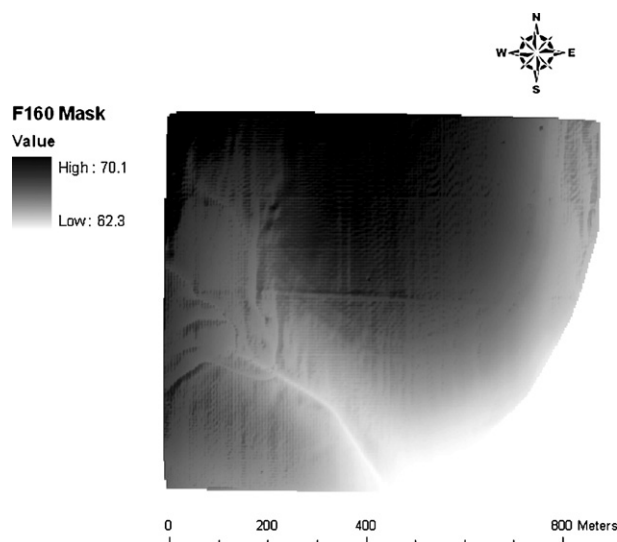
To be able to use the LIDAR data for the Paul Good Farm, the binary data files with multiple flight paths were converted to a mosaic grid format suitable for use with ESRI ARCGIS (ESRI, 1998) and ERDAS Imagine (Leica Geosystems, 2008) software as a digital elevation map. The resulting raster file had a cell size of  $1\text{ m} \times 1\text{ m}$  with a vertical representation of 1 m. The spatial reference system of this DEM was expressed in the geographic coordinate system Geodetic Reference System (GRS) 1980 using the Universal Transverse Mercator (an unprojected coordinate system) with a GRS, 1980 datum (Geodetic Reference Systems 1980, 2008). This spatial reference system can easily be projected onto a planar coordinate system to display data properly or measure distances accurately.

## 6. Yield monitor data

Five yield monitor datasets were used in this analysis. There were three cotton datasets collected in 2001, 2003 and 2004, and two corn datasets collected in 2002 and 2005. The cotton yield datasets were collected using an AgLeader Cotton Yield Monitor model PF 3000 paired with GPS mounted on the cotton pickers with GPS signals collected from satellites and from the subscription service provided by Omnistar. An AgLeader Yield Monitor was used on the corn harvesting equipment with GPS to record the corn yields in 2002 and 2005. The yield monitor data were recorded in North American Datum (NAD) 1983, in the Universal Transversal Mercator UTM system, with the area of interest in Zone 16 North, and were exported from the yield monitors as shape files.

## 7. Software

Three software packages were used in this study. The ESRI ARCGIS 9.2 system and the ERDAS Imagine 9.1 system were used for all aspects of image analysis and spatial data analysis. We found that both ARCGIS and Imagine have their own strengths which could be combined to produce the results desired. The Statistical Analysis System (SAS®) 9.1 and SAS® Enterprise by the SAS® Institute (SAS®, 2008) were used for all statistical analyses. Since all of the yield monitor data were in the NAD 1983 UTM Zone 16N projection (Standardization of Coordinate Systems and Datums, 2008) with meters for land measurement distance, we chose to use this as the projection for all studies reported here.



**Fig. 1.** Detailed F160.Mask shown in hillshade relief which is used to extract data from yield monitor data maps where elevation of the mask is in meters above the ellipsoid.

## 8. Good farm field 160

We chose Field 160 of the Good Farm for our study area, because it had interesting and desirable topographical features. These features included elevation changes over 10 m and aspect exposure for  $360^\circ$ . There were also flat terrain features and a drainage area with substantial slope differentiation. The field consists of 160 acres (65 ha), hence its name. The field is gently rolling land with features from 62 m to 70 m of elevation above the ellipsoid. Drainage flows both to the east and to the southwest with a prominent drainage ditch in the southwest quadrant. Field 160 is shown in Fig. 1 as the grey area in the diagram.

## 9. Mask layers

Field 160 DEM was used as the mask layer for extraction of all features for the spatial analysis of the yield data. The farm DEM was first extracted from the overall DEM created from the LIDAR mosaic image using the polygon extract tool in the ARCGIS Spatial Analyst tool set generating the layer Crop DEM. Then the Crop DEM was shaded using the Hillshade tool to better highlight road features. Next the polygon extract tool was used to extract Field 160 which is bounded by roads on all sides generating the layer F160.Mask. Now F160.Mask is a  $1\text{ m} \times 1\text{ m}$  grid with a 1 m elevation feature.

## 10. Inverse distance weighted (IDW), Kriging and Natural Neighbor layers

All yield monitor datasets were converted from point or vector datasets to grids using the ARCGIS Spatial Analyst toolset so that in the analysis every yield data point from each production year could be used along with data points from topographical features. The yield monitor vector data for cotton and corn are shown in the maps in Figs. 2–6. Generally two harvest machines were used in the harvest operation, but only one was equipped with a yield monitor and GPS. In some parts of the field, generally in the western zone, only the harvester with the yield monitor and GPS equipment was used, producing yield data from all rows. As the harvest proceeded to the east, two harvesters were used side by side, one with a yield monitor and the other without. Thus, interpolation methods had to be used to generate a yield map for the entire field. In some of the eastern areas of Field 160 only the harvester without yield

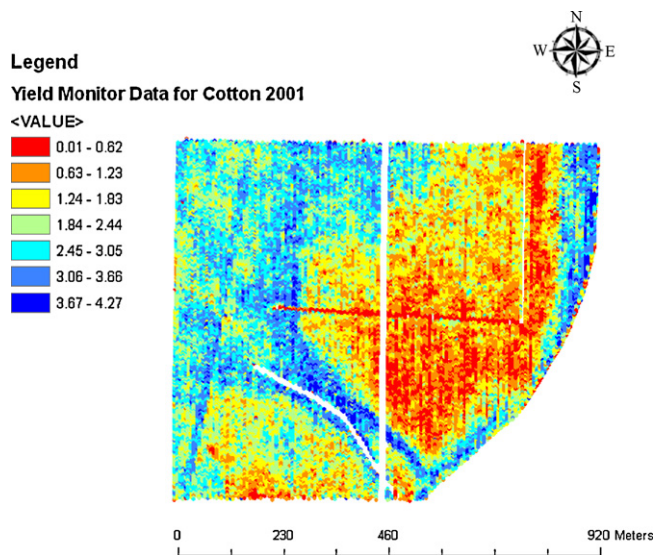


Fig. 2. Yield monitor data map for cotton in 2001.

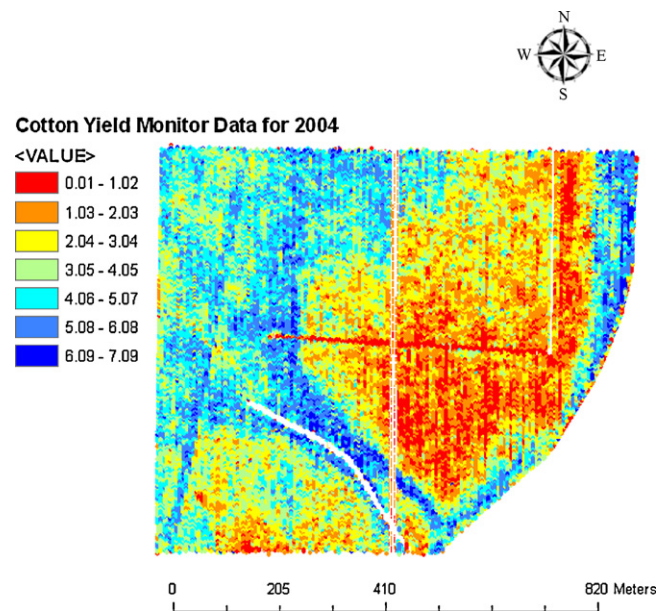


Fig. 5. Yield monitor data map for cotton in 2004.

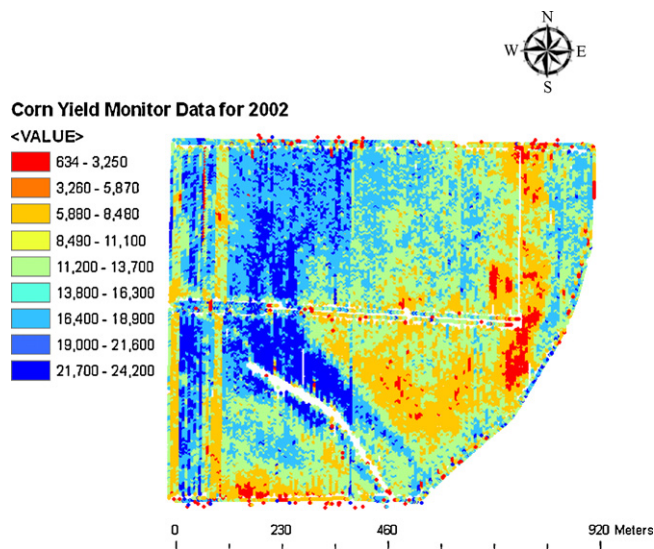


Fig. 3. Yield monitor data map for corn in 2002.

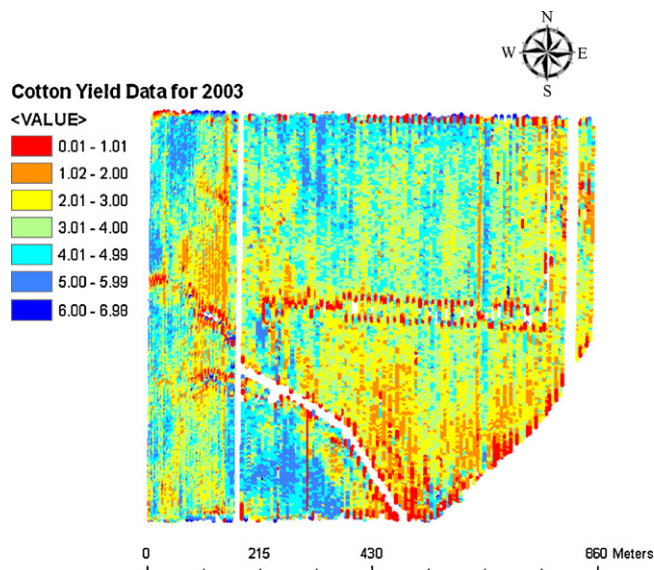


Fig. 4. Yield monitor data map for cotton in 2003.

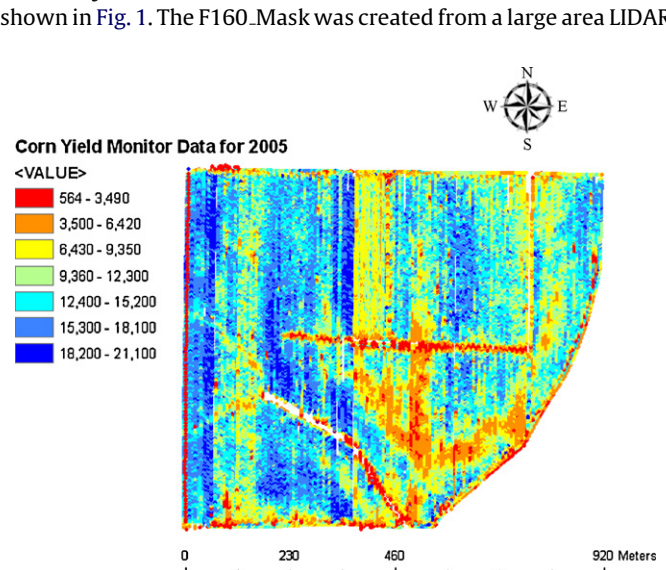


Fig. 6. Yield monitor data map for corn in 2005.

monitoring capability was used, resulting in yield data gaps, and these areas were excluded in the analysis. The tools we used were under the Interpolation heading in ARCGIS Spatial Analyst and were the inverse distance weighted, ordinary point Kriging and Natural Neighbor tools. Given that this process produces interpolated grids, we chose to create a  $3\text{ m} \times 3\text{ m}$  raster grid. A total of fifteen grids were created since there were 5 crop years. Since the Kriging function and the IDW function in ARCGIS create rectangular areas and Field 160 is irregularly shaped, the next step was to reduce the area to only Field 160 area located within the generated rectangle. This was done using the ARCGIS Spatial Analyst tool Extract by Mask. The mask grid layer was F160.Mask. The three sets of grids for IDW, Kriging, and nearest neighbor for the five crop years were to be used for comparison of the interpolation method using simple regression.

## 11. Elevation, aspect, slope, and curvature layers

The layer F160.Mask contains the elevation data for Field 160 as shown in Fig. 1. The F160.Mask was created from a large area LIDAR

map which encompassed the entire Good Farm and surrounding area. The root mean square error of the LIDAR map was 10.7 cm. From this map, a one meter DEM grid was created. Even though LIDAR is very accurate, some artifacts were found in the DEM. After the F160.Mask was created, there were some imperfections found which could affect further analyses. These are called sinks, or errors in data. We used the ARCGIS tool called Fill which is found under the Hydrology heading in Spatial Analyst. The Fill tool removed all of the sinks in the F160.Mask grid. To further smooth the data in the F160.Mask, the ARCGIS tool Block Statistics was used with filtering parameters selected for an annulus filter with an inner radius of 1 m and an outer radius of 3 m with the mean value to be calculated. This created a raster grid layer with smoothed features called F160.Msk.Fcl which was used subsequently to produce all other surface features. Using the tools under Spatial Analyst heading Surface, we next created grids with feature data of aspect, slope and curvature using the respective tools named, appropriately, Aspect, Slope and Curvature.

## 12. Analysis of results and discussion

In all of the regression analyses performed in this study, yield was used as the dependent variable and topographical features of elevation, aspect, curvature and slope derived from a very high density LIDAR image of the field were used as the independent variables. Both simple linear regression and simple quadratic regression were used in the analyses. The LIDAR data were actually degraded from an original image of 0.3 m × 0.3 m pixel resolution. To obtain maps in which the yield data pixel points matched the LIDAR pixel points a number of image processing steps had to be performed which are described subsequently.

Because all of the yield maps and the topographical maps were converted to grids with common geometry, each valid data point (a 3 m × 3 m pixel) could be used in the analysis. This is because each data point for each crop year and each data point for each topographical feature had a common geographic reference point, an Easting and Northing location pair. Even so, there were still data anomalies in the yield maps with zero values from ARCGIS NoData values. These had to be eliminated.

To prepare these above grids for statistical analysis using the SAS® system (SAS, 1985) we used the ERDAS Imagine software. We converted all 15 yield grids and the 4 surface grids into one stack layer consisting of 19 grids along with *x* and *y* data. Since the yield maps had cell sizes of 3 m × 3 m and the surface maps had cell sizes of 1 m × 1 m, the option was chosen in Imagine to create the stack layer using the maximum cell size of 3 m × 3 m for all 19 layers, using the mean for each 9 cells of the surface grids. After manipulating the stack layer to recode missing data to –9999, we converted the resulting stack layer to an ASCII file for import into SAS®.

Using SAS®, all of the data with missing values coded as –9999 from above were eliminated from the dataset. We were then left with a dataset of 19 columns with 65,031 rows. This reduced dataset produced the surface maps shown in Figs. 7–10 which represent elevation, aspect, curvature and slope of the field, respectively with pixels of 3 m × 3 m.

To prevent statistical anomalies from occurring because of scale difference between yield for cotton and yield for corn and to try to equate data for year to year analyses, all of the yield data were normalized using Eq. (1)

$$Y_N = \frac{Y_i - GMAX}{GMAX - GMIN} \times 255 \quad (1)$$

where  $Y_N$  is the normalized result,  $Y_i$  is the yield for each grid point in terms of either IDW, Kriging, or Natural Neighbor interpolation for each cotton or corn crop in years 2001 to 2005, GMAX is the

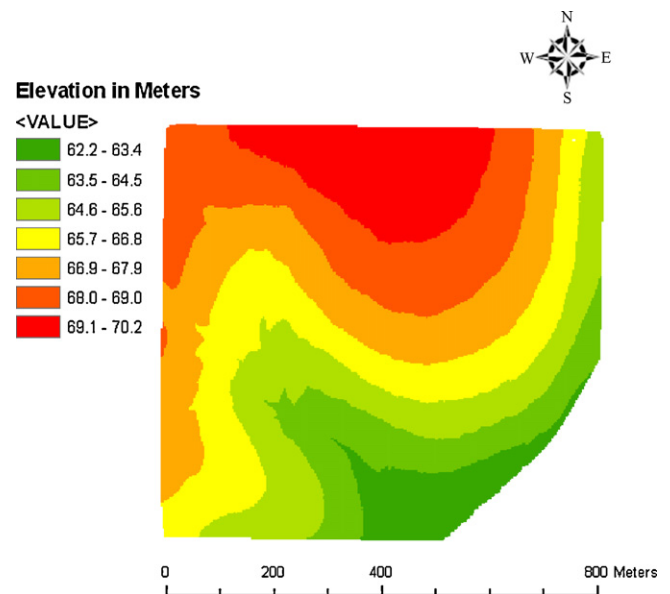


Fig. 7. Elevation surface map for Field 160 in meters above mean sea level.

global maximum for each set and GMIN is the global minimum for each set. The resulting yield data produced by Eq. (1) produces data which is scaled from 0.0 to 255.0. The resulting dataset (a .dbf file) was the final dataset with which all further statistical analyses were performed.

Simple linear regression and simple quadratic regression was performed upon each of the fifteen yield grids using yield as the dependent variable and elevation, aspect, curvature and slope as the independent variables. Results for simple linear regression are given in Table 3. Where as each regression showed that there were statistically significant relationships present, the  $R^2$  values were very low ranging from 0.1% to 11%. In each regression, curvature was nonsignificant with the exception of the regression for the corn data from year 2005 using the Natural Neighbor interpolation method. Simple quadratic regression did not show any significant improvement in explanation and thus is not presented.

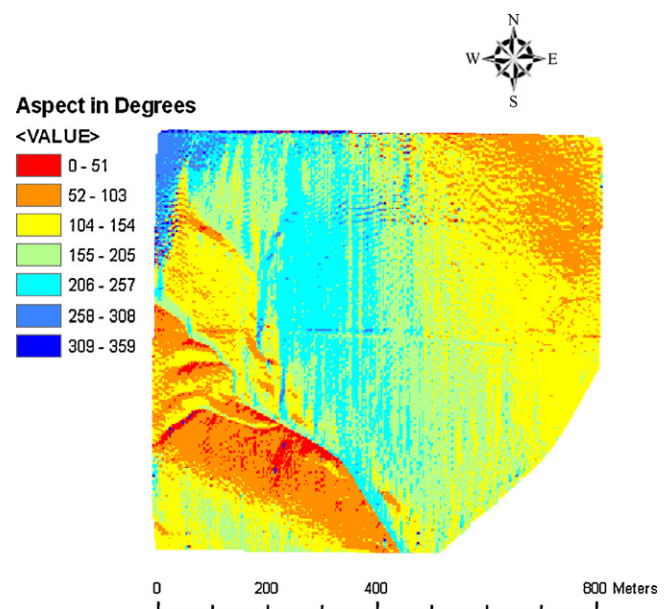
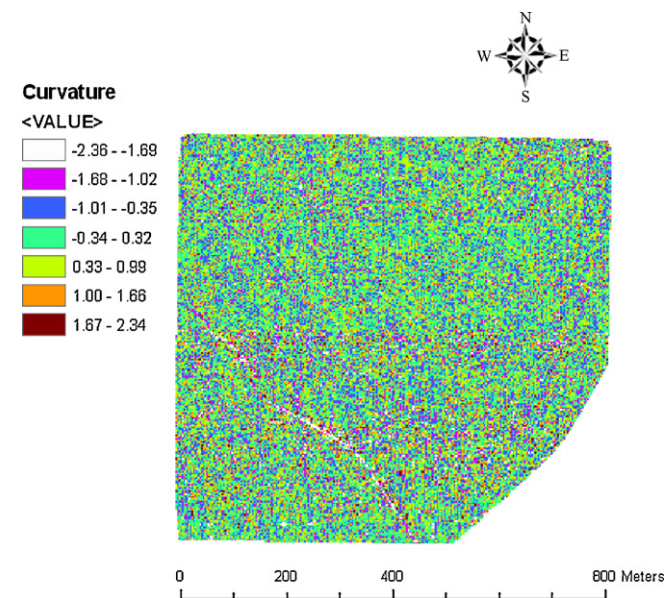


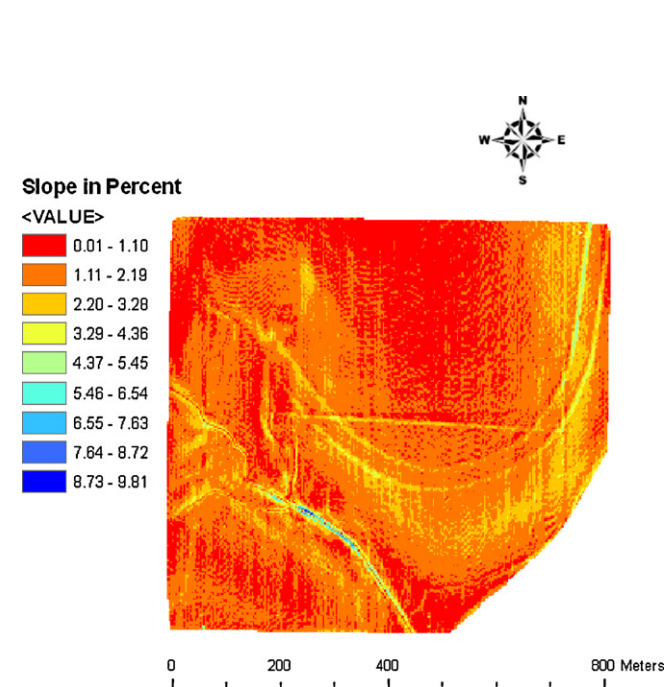
Fig. 8. Aspect surface map for Field 160 in degrees from due north equal zero degrees.

**Table 3**  
Model  $R^2$  values for the fifteen regression datasets generated by five crop years and three interpolation methods and significant independent variables where EL is elevation, AS is aspect, CV is curvature, and SL is slope.

| Regression of the form $Y = I + C_1 \times EL + C_2 \times AS + C_3 \times CV + C_4 \times SL$ |        |                                    |        |                                    |                  |                                    |
|--|--------|------------------------------------|--------|------------------------------------|------------------|------------------------------------|
| Interpolation method   | IDW    |                                    | Kriged |                                    | Natural Neighbor |                                    |
| Statistics   | $R^2$  | Significant at $Pr >  t  < 0.0001$ | $R^2$  | Significant at $Pr >  t  < 0.0001$ | $R^2$            | Significant at $Pr >  t  < 0.0001$ |
| 2001 Cotton  | 0.0468 | EL, AS, SL                         | 0.0925 | EL, SL                             | 0.0659           | EL, SL                             |
| 2002 Corn  | 0.0014 | EL                                 | 0.0987 | EL, AS, SL                         | 0.0500           | EL, AS, EL                         |
| 2003 Cotton  | 0.0949 | EL, AS, SL                         | 0.0659 | EL, SL                             | 0.0014           | EL                                 |
| 2004 Cotton  | 0.1056 | EL, SL                             | 0.0012 | EL                                 | 0.0960           | EL, AS, SL                         |
| 2005 Corn  | 0.0659 | EL, SL                             | 0.0052 | EL, AS                             | 0.1008           | ALL                                |

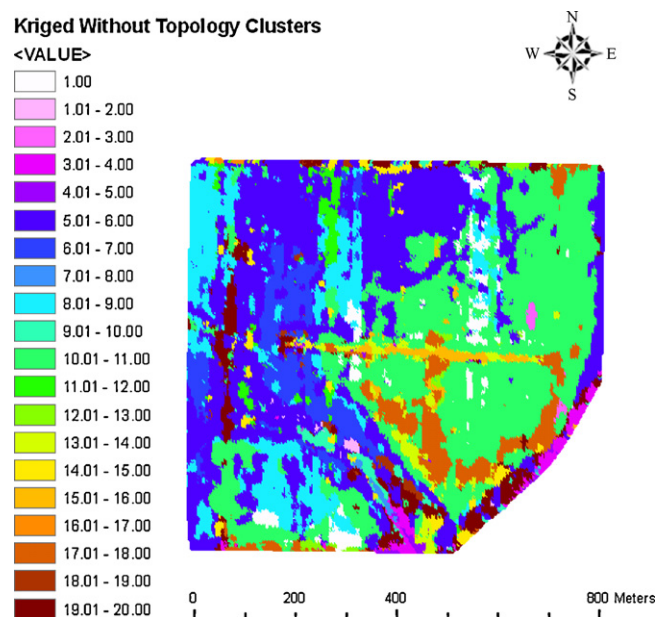


**Fig. 9.** Curvature surface map for Field 160 where a positive number indicates concave down and a negative number indicates convex upward and zero indicates a flat surface.

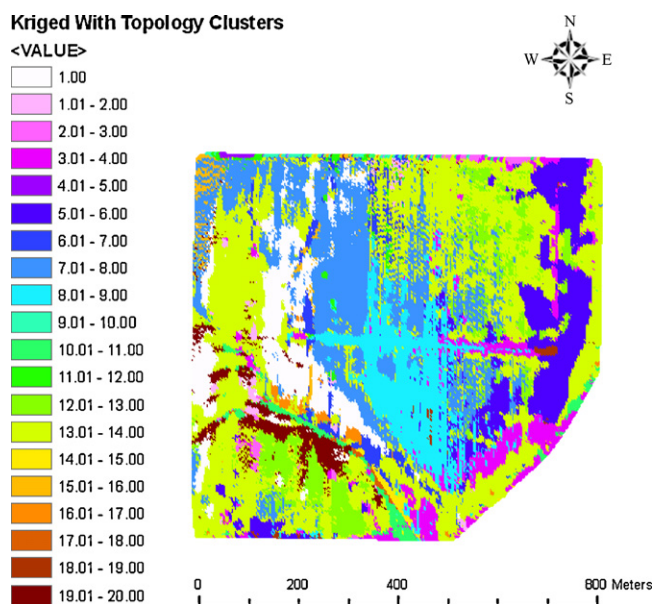


**Fig. 10.** Slope surface map for Field 160 where slope is in per cent with zero being flat and 90% being vertical.

Since we did not find much of the variability explained by simple linear regression analysis for the entire dataset, we reanalyzed the data using the SAS® Fast Cluster algorithm. The entire dataset was used in analyses, producing a new column of data added for each interpolation method both with and without using the four surface features (elevation, aspect, curvature, and slope). These analyses created six new datasets with cluster information of IDW without topology, IDW with topology, Kriging without topology, Kriging with topology, Natural Neighbor without topology and Natural Neighbor with topology. The clustering was performed using yield data. The reason for using clustering was to aggregate similar yield data points together with the express purpose to analyze the yield data points in each cluster for statistical significance. The clusters of yield data points would also represent yield areas which could then be mapped using Geographic Information System (GIS) processing. An observer could then use this yield stability map to relate to topographical differences in the field to yield differences (Cox and Gerard, 2007). These datasets were then imported into ARCGIS and maps were created displaying the cluster results for each interpolation method for the normalized crop yield data. The cluster data for the Kriged data sets are shown in Figs. 11 and 12. In each figure, there are 20 clusters indicated by the legend on the left. The same color scheme is used in each. The value of 20 clusters was chosen so that fine details of transition of field topography could be discerned by the observer.



**Fig. 11.** ARCGIS map of cluster analysis using the Kriging algorithm without topology information and with all five years of scaled crop data. The Fast Cluster algorithm was used in SAS® and data was exported as .dbf file and imported into ARCMAP, converted to a shape file, and displayed using the color legend according to the cluster number.



**Fig. 12.** ARCGIS map of cluster analysis using the Kriging algorithm with topology information and all five years of scaled crop data. The Fast Cluster algorithm was used in SAS® and data was exported as .dbf file and imported into ARCMAP, converted to a shape file, and displayed using the color legend according to the cluster number.

Examining the pair of figures produced by using the Kriging interpolation method both with and without using surface features during the clustering procedure sheds little common information. Each map is very different from the other. One thing that does present itself is the presence of linear features in the maps. We have found that often linear (or straight line) features represent something that the grower has done to the crop (such as planting date or fertilizer changes) and is not a natural physical or biological phenomenon.

Taking the analysis further, each of the six cluster datasets were used in quadratic regressions using the SAS® Proc GLM with the model taking the form:

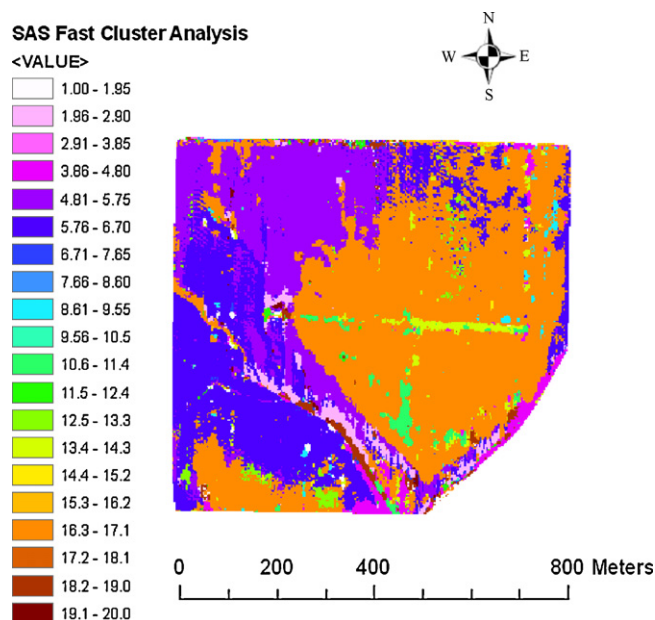
$$Y = \text{Intercept} + X_1 + X_2 + X_3 + X_4 + X_1^2 + X_2^2 + X_3^2 + X_4^2, \quad (2)$$

where  $X_1$  is elevation,  $X_2$  is aspect,  $X_3$  is curvature and  $X_4$  is slope and the dependent  $Y$  values are the yields which are normalized point data from IDW, Kriging, and Natural Neighbor interpolations. Regressions were performed using the results from the clustered datasets both with and without using surface topographical features to produce the clusters. Data were first sorted in SAS® using Easting and Northing, and then sorted again using Cluster. In Proc GLM, the regressions were produced using the 'by cluster' statement.

Results of the regression analyses are given in Tables 4–6 by IDW, Kriging, and Natural Neighbor interpolations, respectively. Inspection of the regression results shows a marked improvement in the  $R^2$  fit. Each cluster analysis with and without surface features produced differing numbers of yield points for each type of cluster. That is, each cluster without surface features for each year had the same number of data points and was different from each cluster with surface features, which then had the same number of yields points for each year, for IDW, Kriging, and Natural Neighbor. In all three tables, there were clusters of yield points which were too few in number to perform regression upon. These clusters were listed as NA in the tables. There were also clusters with enough data points that regression could be performed, but since the number was small, the regression  $R^2$  was unusually high and should probably be disregarded. For each cluster analysis, there were 65,031 yield data points to be allocated to clusters.

Using the above criteria, the regressions in each table show that Kriging, in general, produced a better fit than either IDW or Natural Neighbor interpolations, with Natural Neighbor better than IDW. There was more of a preponderance of higher  $R^2$  for Kriging than for either of the other two interpolation methods. A second observation is that when the surface features are included in the clustering procedure, the  $R^2$  increases for each type of interpolation method. This is very reasonable for one is using more information about the crop and the area in which it grows. This also says that topography and geo-processing can add useful information to the system.

Finally, all fifteen normalized yield vectors using all three interpolation methods were used as analysis variables, and the surface variables were used as grouping variables. Fast Cluster analysis, in Fig. 13, showed that there were 20 statistically important clusters of data over the five year period, regardless of crop planted since the regression analysis showed that for each interpolation method used on the clustered yield data points there was a significant increase in the  $R^2$  value. The areas designated by the colors Light Magenta, Dark Magenta, and Blue represent the primary productivity areas of Field 160. These areas are typified by lower elevation, thereby receiving more water by drainage from the higher elevations. Since the Good Farm uses dryland farming practices, no irrigation water is used. The area of the field represented by the Orange color is the higher elevation areas and is exposed by slope to more direct sunlight. Water drains away from these areas due to elevation and slope, therefore, less water is available to the crops. Yields from the orange area are typically the lowest for Field 160. The objective of this research was to demonstrate that a yield stability map for a commercial row-crop field could be obtained using high density, highly accurate LIDAR imaging as a base map from which topo-



**Fig. 13.** Yield stability map derived from scaling cotton and corn yields to the same dimensionless range of values from a minimum of 0 to a maximum of 255 using the SAS® Proc Fast Cluster algorithm generating 20 data clusters as indicated by the colors of the map for the entire five years of yield data using IDW, Kriging, and Natural Neighbor interpolations and the variables elevation, aspect, curvature and slope. The largest areas of dark and light magentas and blue show clusters where the most productivity from Field 160 occurred for the five years. This area can be described as being in the lower elevation zone with water from rainfall draining into this area from the upper elevations and having either an eastern or western aspect. The large cluster indicated by orange color shows where lower productivity occurred. This area can be described as being on the high elevation zone and with a southern exposure generally having less water availability and more direct sunlight. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

**Table 4**

Regression analysis shown for five crop years of data using IDW interpolation with cluster analysis. Cluster analysis was performed first using only the yield data and then with the surface features of elevation (EL), aspect (AS), curvature (CV), and slope (SL). The regression equation was of the form  $Y = \text{Intercept} + EL + AS + CV + SL + EL^2 + AS^2 + CV^2 + SL^2$ . NA means too few data points for regression. In the VARIABR-Square columns the  $X$  indicates the year the crop was harvested, i.e. 2001 for .1. Unless otherwise indicated by symbols described below, the level of significance for each  $R^2$  value is  $P < 0.0001$ .

| Cluster analysis without surface features included |              |             |             |             |             |             | Cluster analysis without surface features included |             |             |             |             |             |
|--|--------------|-------------|-------------|-------------|-------------|-------------|--|-------------|-------------|-------------|-------------|-------------|
| Cluster number                                     | Observations | R-squared.1 | R-squared.2 | R-squared.3 | R-squared.4 | R-squared.5 | Observations                                       | R-squared.1 | R-squared.2 | R-squared.3 | R-squared.4 | R-squared.5 |
| 1  | 67           | 0.222       | 0.14***     | 0.485       | 0.439       | 0.456       | 536  | 0.659       | 0.68        | 0.53        | 0.37        | 0.263       |
| 2  | 617          | 0.158       | 0.145       | 0.112       | 0.199       | 0.067       | 8  | NA          | NA          | NA          | NA          | NA          |
| 3  | 16039        | 0.057       | 0.07        | 0.056       | 0.137       | 0.032       | 5  | NA          | NA          | NA          | NA          | NA          |
| 4  | 1096         | 0.28        | 0.142       | 0.221       | 0.092       | 0.033       | 245  | 0.27        | 0.179       | 0.177       | 0.332       | 0.027***    |
| 5  | 10           | 0.972***    | 0.767***    | 0.995***    | 0.994***    | 0.99***     | 25   | 0.649**     | 0.585**     | 0.5***      | 0.314***    | 0.534**     |
| 6  | 19           | 0.826*      | 0.937       | 0.638***    | 0.622***    | 0.622***    | 1012   | 0.115       | 0.274       | 0.214       | 0.151       | 0.187       |
| 7  | 5049         | 0.09        | 0.141       | 0.186       | 0.174       | 0.115       | 2  | NA          | NA          | NA          | NA          | NA          |
| 8  | 53           | 0.047***    | 0.369*      | 0.172**     | 0.23***     | 0.38*       | 921  | 0.213       | 0.117       | 0.443       | 0.361       | 0.179       |
| 9  | 161          | 0.195       | 0.416       | 0.54***     | 0.258       | 0.129*      | 5848   | 0.281       | 0.302       | 0.284       | 0.104       | 0.277       |
| 10   | 1414         | 0.263       | 0.52        | 0.298       | 0.185       | 0.102       | 2  | NA          | NA          | NA          | NA          | NA          |
| 11   | 1            | NA          | NA          | NA          | NA          | NA          | 22337  | 0.128       | 0.162       | 0.177       | 0.119       | 0.158       |
| 12   | 233          | 0.387       | 0.119*      | 0.295       | 0.141       | 0.59        | 51   | 0.527       | 0.375*      | 0.187***    | 0.491*      | 0.456*      |
| 13   | 5            | NA          | NA          | NA          | NA          | NA          | 37   | 0.057***    | 0.521       | 0.215***    | 0.187***    | 0.101***    |
| 14   | 31           | 0.55**      | 0.546**     | 0.484**     | 0.618*      | 0.524**     | 3012   | 0.093       | 0.112       | 0.224       | 0.114       | 0.185       |
| 15   | 2791         | 0.105       | 0.193       | 0.2         | 0.067       | 0.045       | 7024   | 0.139       | 0.144       | 0.074       | 0.145       | 0.266       |
| 16   | 26842        | 0.071       | 0.042       | 0.173       | 0.073       | 0.046       | 4167   | 0.327       | 0.226       | 0.341       | 0.031       | 0.049       |
| 17   | 9653         | 0.272       | 0.093.      | 0.189       | 0.089       | 0.099       | 301  | 0.256       | 0.035***    | 0.062**     | 0.42        | 0.373       |
| 18   | 928          | 0.046       | 0.226       | 0.113       | 0.166       | 0.137       | 780  | 0.148       | 0.161       | 0.203       | 0.256       | 0.135       |
| 19   | 18           | 0.627***    | 0.758**     | 0.593**     | 0.405***    | 0.694**     | 22   | 0.443***    | 0.297***    | 0.325***    | 0.368***    | 0.663**     |
| 20   | 4            | NA          | NA          | NA          | NA          | NA          | 18696  | 0.423       | 0.285       | 0.31        | 0.238       | 0.1         |

\* The significance is  $0.01 > P \geq 0.0001$ .

\*\* The level of significance is  $0.1 > P \geq 0.01$ .

\*\*\* The level of significance is  $0.99 > P \geq 0.1$ .

**Table 5**

Regression analysis shown for five crop years of data using Kriging interpolation with cluster analysis. Cluster analysis was performed first using only the yield data and then with the surface features of elevation (EL), aspect (AS), curvature (CV), and slope (SL). The regression equation was of the form  $Y = \text{Intercept} + EL + AS + CV + SL + EL^2 + AS^2 + CV^2 + SL^2$ . NA means too few data points for regression. In the VARIABR-Square columns the  $_X$  indicates the year the crop was harvested, i.e. 2001 for  $_1$ . Unless otherwise indicated by symbols described below, the level of significance for each  $R^2$  value is  $P < 0.0001$ .

| Cluster analysis without surface features included |              |                      |                      |                      |                      |                      | Cluster analysis with surface features included |                      |                      |                      |                      |                      |
|--|--------------|----------------------|----------------------|----------------------|----------------------|----------------------|---|----------------------|----------------------|----------------------|----------------------|----------------------|
| Cluster number                                     | Observations | $R^2$ -squared. $_1$ | $R^2$ -squared. $_2$ | $R^2$ -squared. $_3$ | $R^2$ -squared. $_4$ | $R^2$ -squared. $_5$ | Observations                                    | $R^2$ -squared. $_1$ | $R^2$ -squared. $_2$ | $R^2$ -squared. $_3$ | $R^2$ -squared. $_4$ | $R^2$ -squared. $_5$ |
| 1  | 1667         | 0.198                | 0.457                | 0.492                | 0.224                | 0.099                | 5425  | 0.223                | 0.3                  | 0.245                | 0.054                | 0.166                |
| 2  | 111          | 0.463                | 0.549                | 0.155**              | 0.153**              | 0.23*                | 495   | 0.064                | 0.307                | 0.276                | 0.238                | 0.127                |
| 3  | 154          | 0.556                | 0.136*               | 0.803                | 0.418                | 0.73                 | 307   | 0.137                | 0.468                | 0.249                | 0.515                | 0.382                |
| 4  | 666          | 0.117                | 0.141                | 0.48                 | 0.069                | 0.193                | 3107  | 0.225                | 0.148                | 0.253                | 0.111                | 0.231                |
| 5  | 182          | 0.289                | 0.059***             | 0.614                | 0.302                | 0.275                | 44  | 0.183***             | 0.682                | 0.09***              | 0.466*               | 0.374**              |
| 6  | 19812        | 0.041                | 0.07                 | 0.215                | 0.141                | 0.045                | 6153  | 0.151                | 0.163                | 0.134                | 0.266                | 0.071                |
| 7  | 3941         | 0.155                | 0.264                | 0.254                | 0.088                | 0.149                | 1499  | 0.368                | 0.284                | 0.606                | 0.371                | 0.35                 |
| 8  | 143          | 0.331                | 0.291                | 0.58                 | 0.103**              | 0.127*               | 11106   | 0.447                | 0.237                | 0.242                | 0.219                | 0.061                |
| 9  | 6915         | 0.106                | 0.126                | 0.068                | 0.147                | 0.062                | 6059  | 0.178                | 0.333                | 0.312                | 0.219                | 0.092                |
| 10   | 618          | 0.168                | 0.262                | 0.085                | 0.22                 | 0.244                | 90  | 0.276                | 0.132***             | 0.462                | 0.453                | 0.203**              |
| 11   | 20239        | 0.239                | 0.313                | 0.146                | 0.157                | 0.144                | 889   | 0.237                | 0.143                | 0.416                | 0.492                | 0.196                |
| 12   | 1110         | 0.167                | 0.346                | 0.301                | 0.14                 | 0.085                | 348   | 0.655                | 0.09                 | 0.155                | 0.174                | 0.284                |
| 13   | 419          | 0.509                | 0.36                 | 0.365                | 0.388                | 0.205                | 4359  | 0.527                | 0.348                | 0.264                | 0.347                | 0.172                |
| 14   | 1081         | 0.352                | 0.261                | 0.354                | 0.124                | 0.047                | 21764   | 0.107                | 0.077                | 0.217                | 0.138                | 0.141                |
| 15   | 977          | 0.16                 | 0.213                | 0.328                | 0.06                 | 0.077                | 2312  | 0.21                 | 0.222                | 0.261                | 0.213                | 0.269                |
| 16   | 707          | 0.127                | 0.175                | 0.213                | 0.177                | 0.064                | 663   | 0.407                | 0.473                | 0.357                | 0.268                | 0.164                |
| 17   | 29           | 0.143***             | 0.757                | 0.284***             | 0.328***             | 0.436*               | 351   | 0.401                | 0.444                | 0.168                | 0.249                | 0.103                |
| 18   | 3254         | 0.155                | 0.348                | 0.127                | 0.218                | 0.16                 | 119   | 0.183*               | 0.121**              | 0.632                | 0.399                | 0.583                |
| 19   | 350          | 0.205                | 0.162                | 0.413                | 0.243                | 0.075                | 186   | 0.354                | 0.387                | 0.161                | 0.11*                | 0.202                |
| 20   | 2657         | 0.027                | 0.108                | 0.215                | 0.332                | 0.168                | 1836  | 0.044                | 0.148                | 0.132                | 0.131                | 0.092                |

\* The significance is  $0.01 > P \geq 0.0001$ .

\*\* The level of significance is  $0.1 > P \geq 0.01$ .

\*\*\* The level of significance is  $0.99 > P \geq 0.1$ .

**Table 6**

Regression analysis shown for five crop years of data using Natural Neighbor interpolation with cluster analysis. Cluster analysis was performed first using only the yield data and then with the surface features of elevation (EL), aspect (AS), curvature (CV), and slope (SL). The regression equation was of the form  $Y = \text{Intercept} + EL + AS + CV + SL + EL^2 + AS^2 + CV^2 + SL^2$ . NA means too few data points for regression. In the VARIABR-Square columns the  $X$  indicates the year the crop was harvested, i.e. 2001 for .1. Unless otherwise indicated by symbols described below, the level of significance for each  $R^2$  value is  $P < 0.0001$ .

| Cluster analysis with surface features excluded |              |             |             |             |             |             | Cluster analysis with surface features included |             |             |             |             |             |
|---|--------------|-------------|-------------|-------------|-------------|-------------|---|-------------|-------------|-------------|-------------|-------------|
| Cluster number                                  | Observations | R-squared.1 | R-squared.2 | R-squared.3 | R-squared.4 | R-squared.5 | Observations                                    | R-squared.1 | R-squared.2 | R-squared.3 | R-squared.4 | R-squared.5 |
| 1   | 31           | 0.837       | 0.704*      | 0.258***    | 0.216***    | 0.928       | 8   | NA          | NA          | NA          | NA          | NA          |
| 2   | 5220         | 0.062       | 0.213       | 0.069       | 0.212       | 0.131       | 16612   | 0.069       | 0.104       | 0.083       | 0.128       | 0.215       |
| 3   | 89           | 0.354       | 0.183**     | 0.29*       | 0.369       | 0.104***    | 499   | 0.547       | 0.284       | 0.242       | 0.333       | 0.456       |
| 4   | 1            | NA          | NA          | NA          | NA          | NA          | 1617  | 0.389       | 0.315       | 0.302       | 0.063       | 0.338       |
| 5   | 151          | 0.173*      | 0.202       | 0.195       | 0.03***     | 0.285       | 18  | 0.939       | 0.678**     | 0.576***    | 0.423***    | 0.635**     |
| 6   | 4            | NA          | NA          | NA          | NA          | NA          | 9483  | 0.232**     | 0.367       | 0.31        | 0.149       | 0.182       |
| 7   | 979          | 0.149       | 0.129       | 0.099       | 0.142       | 0.023*      | 133   | 0.327       | 0.618       | 0.455       | 0.263       | 0.56        |
| 8   | 29898        | 0.192       | 0.153       | 0.085       | 0.117       | 0.052       | 12192   | 0.064       | 0.13        | 0.261       | 0.117       | 0.243       |
| 9   | 25           | 0.386***    | 0.247***    | 0.493***    | 0.426***    | 0.631**     | 44  | 0.556       | 0.194***    | 0.386*      | 0.428*      | 0.312**     |
| 10  | 246          | 0.09*       | 0.112*      | 0.169       | 0.25        | 0.135       | 3   | NA          | NA          | NA          | NA          | NA          |
| 11  | 194          | 0.495       | 0.589       | 0.435       | 0.45        | 0.453       | 107   | 0.302       | 0.246*      | 0.24*       | 0.162**     | 0.312       |
| 12  | 2097         | 0.186       | 0.218       | 0.216       | 0.152       | 0.171       | 846   | 0.394       | 0.294       | 0.363       | 0.203       | 0.076       |
| 13  | 2            | NA          | NA          | NA          | NA          | NA          | 1624  | 0.221       | 0.061       | 0.176       | 0.185       | 0.251       |
| 14  | 3            | NA          | NA          | NA          | NA          | NA          | 597   | 0.086       | 0.235       | 0.147       | 0.235       | 0.132       |
| 15  | 110          | 0.599       | 0.182*      | 0.188*      | 0.571       | 0.3         | 13517   | 0.1         | 0.033       | 0.223       | 0.08        | 0.156       |
| 16  | 12535        | 0.121       | 0.164       | 0.227       | 0.039       | 0.146       | 7518  | 0.225       | 0.245       | 0.134       | 0.079       | 0.251       |
| 17  | 9            | NA          | NA          | NA          | NA          | NA          | 60  | 0.423*      | 0.31**      | 0.614       | 0.371*      | 0.635       |
| 18  | 8468         | 0.135       | 0.123       | 0.05        | 0.078       | 0.117       | 29  | 0.607*      | 0.228***    | 0.104***    | 0.422**     | 0.524**     |
| 19  | 59           | 0.108***    | 0.415*      | 0.397*      | 0.342*      | 0.258**     | 53  | 0.277**     | 0.379**     | 0.462*      | 0.536       | 0.545       |
| 20  | 5000         | 0.104       | 0.13        | 0.324       | 0.221       | 0.201       | 71  | 0.14**      | 0.347**     | 0.375*      | 0.483       | 0.444       |

\* The significance is  $0.01 > P \geq 0.0001$ .

\*\* The level of significance is  $0.1 > P \geq 0.01$ .

\*\*\* The level of significance is  $0.99 > P \geq 0.1$ .

graphical information could be derived. Fig. 13 demonstrates such a stability map and has been shown to be statistically valid.

Simple quadratic regression was performed on the dataset used to create Fig. 13. The dataset consisted of all fifteen scaled yields produced from the yield monitor but scaled from 0.0 to 255.0. The dataset also has all of the surface features of elevation, aspect, curvature, and slope. Each of the 65,031 points was referenced by an Easting and Northing location pair. The results are given in Table 7.  $R^2$  from Kriged data generally have the highest values. As can be seen from examining Tables 3–7, there is a steady improvement in the  $R^2$  values over the values in Table 3 indicating more accounting for explanatory information as the procedures were refined.

Because the base map for all calculations involving topographical variables used in this analysis were derived from a very accurate, dense LIDAR map of elevation from which a digital elevation map was created, the analysis was carried out without further consideration of spatial structure (Lowenberg-Deboer et al., 2008). We believed that because of the density and accuracy of the LIDAR DEM that we could explore other methods to explain possible relationships of crop yield stability. However, since the Kriged data method proved to have the most explanatory capability of the yield variability in the clusters of those explored, this possibly implies that there is some spatial structure involved with these datasets. We propose to explore this possibility in future work. While the number of clusters was preset at 20 in this analysis, this clustering level was chosen so that fine levels of transition from one yield area to another could be seen in the resulting GIS map display. The regression analysis performed on the yield data points associated with each cluster showed a steady increase in the  $R^2$  values as more information was added into the regressions. Each time more information was added, the resulting cluster analysis showed that differing sets of yield data points would become associated with the new cluster sets of yield data points. We propose to add further information to this analysis in future work and to reduce the number of clusters to a number which growers could more reliably interpret and use.

**Table 7**

Regression analysis shown for the combined dataset using Cluster Analysis for all fifteen scaled yield data points and the four surface features. The regression equation was of the form  $Y = \text{Intercept} + EL + AS + CV + SL + EL^2 + AS^2 + CV^2 + SL^2$ . NA means too few data points for regression. There are three  $R^2$  values for each of the 20 cluster values, one each for IDW, Kriged, and Natural Neighbor. All  $R^2$  values are significant at  $P < 0.05$  unless denoted otherwise by the symbols described subsequently.

| Regression equation $Y = \text{Int} + EL + AS + CV + SL + EL^2 + AS^2 + CV^2 + SL^2$ |              |                    |                     |                           |
|--|--------------|--------------------|---------------------|---------------------------|
| Cluster number   | Observations | $R^2$ Square IDW % | $R^2$ Square Krig % | $R^2$ Square Nat Neigh. % |
| 1  | 313          | 28                 | 35                  | 21                        |
| 2  | 1512         | 7                  | 9                   | 7                         |
| 3  | 30           | 22*                | 58*                 | 49*                       |
| 4  | 1556         | 22                 | 30                  | 22                        |
| 5  | 12261        | 9                  | 11                  | 8                         |
| 6  | 15020        | 14                 | 16                  | 13                        |
| 7  | 184          | 9                  | 12                  | 6                         |
| 8  | 40           | 27                 | 32*                 | 25*                       |
| 9  | 363          | 21                 | 23                  | 19                        |
| 10   | 24           | 42**               | 32**                | 39**                      |
| 11   | 870          | 28                 | 40                  | 28                        |
| 12   | 184          | 27                 | 33                  | 20                        |
| 13   | 675          | 27                 | 32                  | 23                        |
| 14   | 1021         | 24                 | 31                  | 27                        |
| 15   | 91           | 24                 | 29                  | 18                        |
| 16   | 89           | 28*                | 40                  | 17***                     |
| 17   | 29821        | 23                 | 29                  | 22                        |
| 18   | 37           | 34                 | 45                  | 52                        |
| 19   | 804          | 11                 | 12                  | 10                        |
| 20   | 136          | 22****             | 29                  | 24                        |

\* The  $R^2$  value means significance at  $0.9 < P < 0.1$ .

\*\* The  $R^2$  value means significance at  $0.8 < P < 0.1$ .

\*\*\* The  $R^2$  value means significance at  $0.09 < P < 0.1$ .

\*\*\*\* The  $R^2$  value means significance at  $0.6 < P < 0.06$ .

### 13. Conclusions and further research

- Simple quadratic regression analysis of the yield data points that produced these maps indicated that Kriging produced higher explanatory statistics than either IDW or Natural Neighbor interpolation methods. When the entire dataset, using IDW, Kriging, and Natural Neighbor interpolations of scaled yields combined with surface features of elevation, aspect, curvature and slope, was used with the SAS® Fast Cluster Analysis, this procedure provided the most information about Field 160 over the five year period.
- Lower elevations in the field represented the most productive areas while the higher elevations exposed to the southern aspect were found to be the least productive over the five year study period. Potential exists for the grower to focus more resources upon the more productive areas to maintain higher productivity, especially in years where less rainfall is predicted.
- The objective of this research was to show that a statistically valid yield stability map for multi-crops grown over several years could be developed for a commercial field using LIDAR imaging as the base map. This was accomplished.
- A Class A weather station which collects temperature, daily wind run, rainfall, dew point and solar radiation is located on the farm, and weather data for all five years is available, representing potential variables for further analysis of the crop datasets. In addition, electrical conductivity, Veris data, for the soil is available and will be used for additional analyses. Crop yield data for years 2006, 2007, and 2008 are also being available and will also be used in further analyses. The question of spatial structure as indicated by the Kriging analysis will also be examined.

### References

- Ackerman, F., 1999. Airborne laser scanning—present status and future expectations. *ISPRS Journal of Photogrammetry and Remote Sensing* 54, 64–67.
- Axelsson, P., 1999. Processing of laser scanning data—algorithms and applications. *ISPRS Journal of Photogrammetry Remote Sensing* 54, 138–147.
- Baltsavias, E.P., 1999. Airborne laser scanning: basic relations and formulas. *ISPRS Journal of Photogrammetry Remote Sensing* 54, 199–214.
- Barnes, F.J., Karl, R.J., Kunkel, K.E., Stone, G.L., 1990. Lidar determination of horizontal and vertical variability in water vapor over cotton. *Remote Sensing of Environment* 32, 81–90.
- Baumgardner, M.F., Silva, L.F., Biehl, L.L., Stoner, E.R., 1995. Reflectance properties in soils. *Advances in Agronomy* 38, 1–44.
- Ben-Dor, E., Del Barrio, G., Eperma, G., 1998. Soil reflectance. *Remote Sensing Reviews* 13, 300–306.
- Bullock, D.S., Bullock, D.G., 2000. From agronomic research to farm management guidelines: a primer on the economics of information and precision technology. *Precision Agriculture* 2, 71–101.
- Bullock, D.S., Kitchen, N., Bullock, D.G., 2007. Multidisciplinary teams: a necessity for research in precision agriculture systems. *Crop Science* 47, 1765–1769.
- Changere, A., Lal, R., 1997. Slope position and erosional effects on soil properties and corn production on a Miamian soil in Central Ohio. *Journal of Sustainable Agriculture* 11, 5–21.
- Colvin, T.S., Jaynes, D.B., Karlen, D.L., Laird, D.A., Ambuel, J.R., 1997. Yield variability within a central Iowa field. *Transactions of the American Society of Agricultural Engineers* 40, 883–889.
- Cox, M.S., Gerard, P.D., 2007. Soil management zone determination by yield stability analysis and classification. *Agronomy Journal* 99, 1357–1365.
- Drysdale (Warren), G.M.R., Metternicht, G.I., 2003. Utilising remote sensing and terrain data for designing multi-scale sampling strategies of soil properties in agricultural fields. *Proceedings of Spatial Sciences 2003*, Canberra, ACT, September, CD Rom.
- ESRI, 1998. *ESRI Shapefile Technical Description*. ESRI, Redlands, CA.
- Fleming, K.L., Westfall, D.G., Wiens, D.W., Brodhall, M.C., 2000. Evaluating farmer defined management zones for variable rate fertilizer applications. *Precision Agriculture* 2, 201–215.
- Filin, S., 2004. Surface classification from airborne laser scanning data. *Computers & Geosciences* 30, 1033–1041.
- Frtich, D., Kilian, J., 1994. Filtering and calibration of laser scanner measurements. *International Archives of Photogrammetry and Remote Sensing* 30 (3), 227–234.
- GIS Dictionary, 2009. <http://support.esri.com/index.cfm?fa=knowledgebase.gisDictionary.search&search=true&searchTerm=dem> (Verified 28 Sept., 2009).
- Geodetic Reference Systems, 1980. 2008. [http://en.wikipedia.org/wiki/Geodesy#Geoid\\_and\\_reference\\_ellipsoid](http://en.wikipedia.org/wiki/Geodesy#Geoid_and_reference_ellipsoid) (Verified 22 Sept. 2008).

- Hanna, A.Y., Harlan, P.W., Lewis, D.T., 1982. Soil available water as influenced by landscape position and aspect. *Agronomy Journal* 74, 999–1004.
- Hollaus, M., Wagner, W., Kraus, K., 2005. Airborne laser scanning and usefulness for hydrological models. *Advances in Geosciences* 5, 57–63.
- Hornung, A., Khosla, R., Reich, R., Inman, D., Westfall, D.G., 2006. Comparison of site-specific management zones: soil-color-based and yield-based. *Agronomy Journal* 98, 407–415.
- Geodetic Reference System, 1980. <http://en.wikipedia.org/wiki/GRS.80> (Link Tested 8/28/2008).
- Huising, E.J., Pereira, L.M., 1998. Errors and accuracy estimates of laser data acquired by various laser scanning systems for topographic applications. *ISPRS Journal of Photogrammetry and Remote Sensing* 53 (5), 245–261.
- Intermap Technologies, 2009. <https://www.vtrentz.net/imarkownerfiles/ownerassets/868/brochure.DEMs.pdf> (Link Tested 29 Sept., 2009).
- Irons, J.R., Weismiller, R.A., Petersen, G.W., 1989. Soil reflectance. In: Asrar, G. (Ed.), *Theory and Applications of Optical Remote Sensing*. Wiley, New York, USA, pp. 67–106.
- Jaynes, D.B., Colvin, T.S., 1997. Spatiotemporal variability of corn and soybean yield. *Agronomy Journal* 89, 30–37.
- Jensen, J.R., 2000. Active and passive microwave and LiDAR remote sensing. In: *Remote Sensing of the Environment: An Earth Resource Perspective*. Prentice-Hall, New Jersey, pp. 285–332.
- Jones, A.J., Mielke, L.N., Bartles, C.A., Miller, C.A., 1989. Relationship of landscape position and properties to crop production. *Journal of Soil and Water Conservation* 44, 328–332.
- Kaspar, T.C., Colvin, T.S., Jaynes, D.B., Karlen, D.L., James, D.E., Meek, D.W., Pulido, D., Butler, H., 2003. Relationship between six years of corn yields and terrain attributes. *Precision Agriculture* 4, 87–101.
- Kraus, K., Pfeifer, N., 1998. Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry Remote Sensing* 53 (4), 193–203.
- Kravchenko, A., Robertson, G.P., Thelen, K.D., Harwood, R.R., 2005. Management, topographical, and weather effects on spatial variability of crop grain yields. *Agronomy Journal* 97, 514–523.
- Kravchenko, A., Bullock, D.G., 2000. Correlation of corn and soybean grain yield with topography and soil properties. *Agronomy Journal* 92, 75–83.
- Kravchenko, A.N., Bullock, D.G., 2002a. Spatial variability of soybean quality data as a function of field topography: I. Spatial data analysis. *Crop Science* 42, 804–815.
- Kravchenko, A.N., Bullock, D.G., 2002b. Spatial variability of soybean quality data as a function of field topography: II. A proposed technique for calculating the size of the area for differential soybean harvest. *Crop Science* 42, 816–821.
- Lark, R.M., Stafford, J.V., Bolam, H.C., 1997. Limitations on the spatial resolution of yield mapping for combinable crops. *Journal of Agricultural Engineering Research* 66, 183–193.
- Latypov, D., 2002. Estimating relative lidar accuracy information from overlapping flight lines. *ISPRS Journal of Photogrammetry Remote Sensing* 56, 236–245.
- Leica Geosystems, 2008. [http://gi.leica-geosystems.com/documents/pdf/Powering\\_GI\\_Chain.White.Paper.pdf](http://gi.leica-geosystems.com/documents/pdf/Powering_GI_Chain.White.Paper.pdf) (Link Tested 9/4/2008).
- Leyva, R.I., Henry, R.J., Graham, L.A., Hill, J.M., 2002. Use of LIDAR to determine vegetation vertical distribution in areas of potential black-capped vireo habitat at Fort Hood, Texas. *Endangered Species Monitoring and Management at Fort Hood, Texas, 2002 Annual Report*. The Nature Conservancy, Fort Hood, TX.
- Lowenberg-DeBoer, J.M., Griffin, T.W., Florax, R.J.G.M., 2008. Use of cross regression to model local spatial autocorrelation in precision agriculture. <http://www.aragriculture.org/precisionag/research/Autocorrelation.pdf> (Link tested December, 23, 2008).
- Means, J.E., Acker, S.A., Fitt, B.J., Renslow, M., Emerson, L., Hendrix, C.J., 2000. Predicting forest stand characteristics with airborne scanning LiDAR. *Photogrammetric Engineering & Remote Sensing* 54, 95–104.
- Miller, M.P., Singer, M.J., Nielsen, D.R., 1988. Spatial variability of wheat yield and soil properties on complex hills. *Soil Sci. Soc. Am. J.* 52, 1133–1141.
- Moore, I.D., Gessler, P.E., Nielson, G.A., 1993. Soil attribute prediction using terrain analysis. *Soil Science Society of American Journal* 57, 443–452.
- Morin, K.W., 2002. Calibration of airborne laser scanners. UCGE Reports No. 20179, Department of Geomatics Engineering, University of Calgary, Calgary, Alberta, CA [URL: <http://www.geomatics.ucalgary.ca/links/GradTheses.html>] (Verified 20 March 2007).
- Mueller, T.G., Pierce, F.J., Schabengerger, O., Warncke, D.D., 2001. Map quality for site-specific fertility management. *Soil Science Society of American Journal* 65, 1547–1558.
- Popescu, S.C., Wynne, R.H., Nelson, R.F., 2002. Estimating plot-level tree heights with lidar: local filtering with a canopy-height based variable window size. *Computers and Electronics in Agriculture* 37, 71–95.
- SAS®, 2008. SAS® Enterprise Guide. <http://www.sas.com/technologies/bi/query-reporting/guide/index.html> (Link Tested 9/4/2008).
- Shan, J., Sampath, A., 2005. Urban DEM generation from raw lidar data: a labeling algorithm and its performance. *Photogrammetric Engineering and Remote Sensing* 71 (2), 217–226.
- Sithole, G., Vosselman, G., 2004. Experimental comparison of filter algorithms for bare-earth extraction from airborne laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing* 59, 85–101.
- Skalud, J., Lichti, D., 2006. Rigorous approach to bore-sight self-calibration in airborne laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing* 61, 47–49.
- Spomer, R.G., Piess, R.F., 1982. Soil productivity and erosion of Iowa loess soils. *Transactions of the American Society of Agricultural Engineers* 25, 1295–1299.
- Standardization of Coordinate Systems and Datums, <http://wvgs.wvu.edu/otherdocs/standardsandpubs/wv.coordinate.systems.aug02.html> (Link Tested 8/28/2008).
- Stone, J.R., Gilliam, J.W., Cassel, D.K., Daniels, R.B., Nelson, L.A., Kleiss, H.J., 1985. Effect of erosion and landscape position on the productivity of Piedmont soils. *Soil Science Society of America Journal* 49, 987–991.
- Timlin, D.J., Pachepsky, Ya., Snyder, V.A., Bryant, R.B., 1998. Spatial and temporal variability of corn grain yield on a hillslope. *Soil Science Society of America Journal* 62, 764–773.
- Vosselman, G., 2002. On the estimation of planimetric offsets in laser altimetry data. *International Archives of Photogrammetry and Remote Sensing* 34 (3A), 375–380.
- Warrick, A.W., Myers, D.E., Nielsen, D.R., 1986. Geostatistical methods applied to soil science. In: Klute et al. (Eds.) *Methods of Soil Analysis*. Part 1. 2nd ed. Agron. Monogr. 9. ASA and SSSA, Madison, WI, pp. 53–82.
- Wehr, A., Lohr, U., 1999. Airborne laser scanning—an introduction and overview. *ISPRS Journal of Photogrammetry Remote Sensing* 54, 68–82.
- Willers, J.L., Jin, M., Eksioğlu, B., Zusmanis, A., O'Hara, C., Jenkins, J. A post-processing step error correction algorithm for overlapping LiDAR strips from agricultural landscapes. *Computers and Electronics in Agriculture*, in press.